

СТАНОВИЩЕ

по процедура за защита на дисертационен труд
за придобиване на образователна и научна степен „доктор“

от кандидат: **Валентин Венков Змийчаров**

на тема: *“Автоматично резюмиране на правни текстове“*,

Област на висше образование: **4. Природни науки, математика и информатика**

Професионално направление: **4.6. Информатика и компютърни науки**

Докторска програма: **"Софтуерни технологии" – Откриване на знания,**

катедра: **„Софтуерни технологии“**,

Факултет по математика и информатика (ФМИ),

Софийски университет „Св. Климент Охридски“ (СУ).

Становището е изготвено от **проф. д.н. Даниела Ананиева Орозова, Тракийски университет - Стара Загора** в качеството ми на член на научното жури, съгласно Заповед № РД 38-278/ 9.06.2025 г. на Ректора на Софийския университет.

1. Обща характеристика на дисертационния труд и представените материали

Автор на дисертационния труд е Валентин Венков Змийчаров – докторант в задочна форма на обучение към катедра „Софтуерни технологии“ на Факултета по математика и информатика при СУ „Св. Климент Охридски“ с научен ръководител проф. д-р Иван Койчев. Обект на изследване в дисертационния труд е приложението на съвременни методи за обработка на естествен език в контекста на сложни и дълги текстове, като правни документи и научни статии.

Дисертационният труд на Валентин Змийчаров на тема *„Автоматично резюмиране на правни текстове“* е в общ обем от 114 страници. Включва 27 фигури и 9 таблици. Текстът е структуриран в шест глави и използвана литература. Проложени са списъци на фигурите и таблиците, както и декларация за оригиналност на резултатите. Библиографията обхваща 58 източника, всички на английски език.

Целта на изследването е дефинирана на стр. 8 в дисертационния труд: *„Разработване на методология за автоматично резюмиране на дълги правни текстове на различни езици чрез използване на съвременни езикови модели и адаптирани подходи“*. Поставените 5 задачи кореспондират с целта и приложените методи.

В първата глава *„Въведение“* се дава описание на изследваната тема, актуалността на задачата и мотивацията за провеждане на изследването. Дефинирани са целта и задачите в дисертационния труд.

Във втората глава *„Обзор на предметната област“* е направен преглед на основните концепции и алгоритми в областта на обработката на естествен език и автоматичното резюмиране. Направен е анализ на наличните набори от данни за правни текстове, като се акцентира върху тяхната пригодност за обучение на съвременни модели за резюмиране.

В трета глава *„Нов корпус от данни за правни текстове и техните резюмета“* е представен набор от правни текстове, обхващащ 24 езика, като са описани базовите им характеристики. Представен е процеса на събиране, изчистване и подготовка на данните, използвани в експериментите. Описани са проблеми при многоезиковата обработка и са включени сравнения между данните на различни езици.

Четвърта глава *„Разработени и адаптирани подходи за резюмиране на текстове“* представя основните методи и техники за резюмиране на текстове. Разработеният модел, наречен StructExtSum използва структурата на информацията в документите за подобряване на резюметата.

В пета глава *„Резултати от проведените експерименти“* се прави анализ на ефективността на използваните методи, като се акцентира на приложимостта на подходите за специфични езици и типове документи. Дава се оценка на резултатите, постигнати с големи езикови модели (GPT) и сравнението им с други подходи.

Шестата глава *„Заклучение и бъдеща работа“* обобщава постигнатите основни резултати в дисертационния труд и очертава бъдещи насоки на развитие.

2. Данни и лични впечатления за кандидата

Представянето на проблема и направеният преглед на основните концепции и алгоритми в областта на обработката на естествен език, както и направените изводи и насоки за бъдещо развитие на изследванията ми дават основание да заключа, че докторантът познава задълбочено проблема на дисертационния труд.

Не познавам лично Валентин Венков Змийчаров, но получих отличното впечатление от дисертационния труд и значимите резултати, представени в него.

3. Съдържателен анализ на научните и научно-приложните постижения на кандидата, съдържащи се в представения дисертационен труд и публикациите към него, включени по процедурата

Настоящата дисертация предоставя значителни приноси както в теоретичен, така и в практически аспект, свързани с областта на автоматичното резюмиране на текстове. Приемам основните резултати и приноси, посочени от докторанта на страници 96 и 97 в дисертационния труд. Приносите имат научен и научно-приложен характер.

Разработен е иновативен екстрактивен метод *StructExtSum* за автоматично резюмиране на правни текстове, основан на анализ на документната структура. Изследвана е ефективността на съществуващите езикови модели върху множество езици. Направен е сравнителен анализ на предварително обучени модели, който може да бъде разглеждан като насоки за избор на подходящ модел в зависимост от специфичния контекст. Изграден и публично достъпен е многоезиков набор от данни за резюмиране на правни документи, включващ 24 езика. Разработените алгоритми могат да бъдат внедрени в софтуерни платформи за управление на документи или да бъдат интегрирани в правни системи за автоматично извличане на съществени части от текстове.

4. Аprobация на резултатите

Резултатите от дисертационното изследване са представени в 4 научни публикации по темата на дисертационния труд в специализирани научни издания с SJR за годината на публикуването.

Всички публикации са на английски език.

Всички публикацииите са в съавторство, като са представени декларации на съответните авторски колективи за равностоен личен принос.

Публикациите носят общо 120 точки, при необходими 30 точки по минималните национални изисквания.

След подробен преглед на представените научни трудове установих, че:

а) научните трудове отговарят на минималните национални изисквания (по чл. 2б, ал. 2 и 3 на ЗРАСРБ) и съответно на допълнителните изисквания на СУ „Св. Климент Охридски“ за придобиване на образователна и научна степен „доктор“ в научната област и професионално направление на процедурата;

б) представените от кандидата резултати в дисертационния труд и научни трудове към него не повтарят такива от предишни процедури за придобиване на научно звание и академична длъжност;

в) няма доказано по законоустановения ред плагиатство в представения дисертационен труд и научни трудове по тази процедура.

5. Качества на автореферата

Авторефератът (на български и английски език) отговаря по обем и съдържание на изискванията за точно, пълно и сбито отразяване на дисертационния труд.

6. Критични бележки и препоръки

Нямам забележки по представените материали. Документите показват успешна и резултатна изследователска дейност на Валентин Змийчаров. Препоръчвам на кандидата в своите научни трудове да изследва и цитира публикации и на български автори.

7. Заключение

След като се запознах с представените в процедурата дисертационен труд и придружаващите го научни трудове и въз основа на направения анализ на тяхната значимост и съдържащи се в тях научни и научно-приложни приноси, **потвърждавам**, че представеният дисертационен труд и научните публикации към него, както и качеството и оригиналността на представените в тях резултати и постижения, отговарят на изискванията на ЗРАСРБ, Правилника за приложението му и съответния Правилник на СУ „Св. Климент Охридски“ за придобиване от кандидата на образователната и научна степен „доктор“ в научната област 4. Природни науки, математика и информатика и професионално направление: 4.6. Информатика и компютърни науки. В частност кандидатът удовлетворява минималните национални изисквания в професионалното направление и не е установено плагиатство в представените по конкурса научни трудове.

Въз основа на гореизложеното, давам своята **положителна оценка** за проведеното изследване, представено от рецензираните по-горе дисертационен труд, автореферат, постигнати резултати и приноси, и *предлагам на почитаемото научно жури да присъди образователната и научна степен „доктор“ на Валентин Венков Змийчаров* в област на висше образование: **4. Природни науки, математика и информатика**, професионално направление **4.6. Информатика и компютърни науки**, докторска програма **"Софтуерни технологии" - Откриване на знания.**

29.07. 2025 г.

Изготвил становището:

(проф. дн Даниела Орозова)